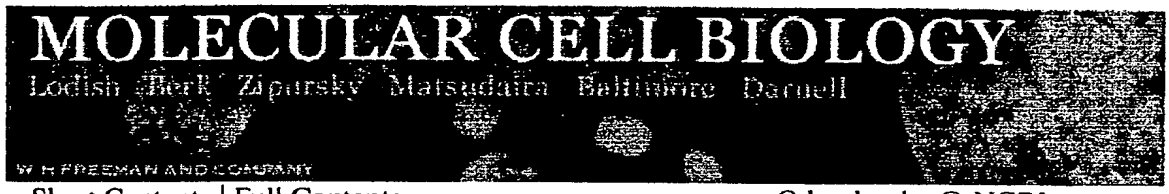
EXHIBIT 21

Amendment in Response to
November 3, 2008 Office Action
Submitted: May 4, 2009

Serial No. 10/821,726

Filed: April 8, 2004

Applicants: Michael Wayne Graham et al.



Short Contents | Full Contents Other books @ NCBI

Molecular Cell Biology ➔ **11. RNA Processing, Nuclear Transport, and Post-Transcriptional Control**

11.2. Processing of Eukaryotic mRNA

As discussed in Chapter 4, the initial primary transcript synthesized by RNA polymerase II undergoes several processing steps before a functional mRNA is produced. In this section, we take a closer look at how eukaryotic cells carry out mRNA processing, which includes three major processes: **5′ capping, 3′ cleavage/polyadenylation,** and RNA splicing (Figure 11-7). Processing occurs in the nucleus, and the functional mRNA produced is transported to the cytoplasm by mechanisms discussed later.

The 5′-Cap Is Added to Nascent RNAs Shortly after Initiation by RNA Polymerase II

After nascent RNA molecules produced by RNA polymerase II reach a length of 25 – 30 nucleotides, *7-methylguanosine* is added to their 5′ end. This initial step in RNA processing is catalyzed by a dimeric capping enzyme, which associates with the phosphorylated carboxyl-terminal tail domain (CTD) of RNA polymerase II. Recall that the CTD becomes phosphorylated during transcription initiation (see Figure 10-50). Because the capping enzyme does not associate with polymerase I or III, capping is specific for transcripts produced by RNA polymerase II.

One subunit of the capping enzyme removes the γ-phosphate from the 5′ end of the nascent RNA emerging from the surface of a RNA polymerase II (Figure 11-8). The other subunit transfers the GMP moiety from GTP to the 5′-diphosphate of the nascent transcript, creating the guanosine 5′-5′-triphosphate structure. In the final steps, separate enzymes transfer methyl groups from *S*-adenosylmethionine to the N_7 position of the guanine and the 2′ oxygens of riboses at the 5′ end of the nascent RNA. ↑ TOP

Pre-mRNAs Are Associated with hnRNP Proteins Containing Conserved RNA-Binding Domains

Nascent RNA transcripts from protein-coding genes and mRNA processing intermediates, collectively referred to as pre-mRNA, do not exist as free RNA molecules in the nuclei of eukaryotic cells. From the time nascent transcripts first emerge from RNA polymerase II until mature mRNAs are transported into the cytoplasm, the RNA molecules are associated with an abundant set of nuclear proteins, as numerous in growing eukaryotic cells as histones. These proteins are the major protein

components of heterogeneous ribonucleoprotein particles (hnRNPs), which contain *heterogeneous nuclear RNA* (hnRNA), a collective term referring to pre-mRNA and other nuclear RNAs of various sizes. The proteins in these ribonucleoprotein particles can be dramatically visualized with fluorescentlabeled monoclonal antibodies (Figure 11-9).

To identify **hnRNP proteins,** researchers exposed cells to high-dose UV irradiation, which causes covalent cross-links to form between RNA bases and closely associated proteins. Chromatography of nuclear extracts from treated cells on an oligo-dT cellulose column, which binds RNAs with a poly(A) tail, was used to recover proteins that had become cross-linked to nuclear mRNA in living cells (i.e., hnRNP proteins). Subsequent treatment of cell extracts from unirradiated human cells with monoclonal antibodies specific for the major hnRNP proteins identified by this cross-linking technique revealed a complex set of abundant hnRNP proteins ranging in size from 34 to 120 kDa. Characterization of the mRNAs encoding these proteins has shown that some of them (e.g., A2 and B1) are related proteins derived by alternative splicing of exons from the same transcription unit.

Binding studies with purified hnRNP proteins suggest that different hnRNP proteins associate with different regions of a newly made pre-mRNA molecule as determined by the sequence of the RNA. For example, the hnRNP proteins A1, C, and D bind preferentially to the pyrimidine-rich sequences at the 3′ ends of introns, discussed in a later section. Like transcription factors, most hnRNP proteins have a modular structure. They contain one or more RNA-binding domains and at least one other domain that is thought to interact with other proteins. Several different RNA-binding motifs have been identified by constructing deletions of hnRNP proteins and testing their ability to bind RNA. Although some RNA-binding proteins contain domains with the zinc-finger motif common in DNA-binding proteins (see Figure 10-41), this motif has not yet been described in any hnRNP proteins.

The *RNP motif,* also called the RNA-binding domain (RBD), is the most common RNA-binding domain in hnRNP proteins. This ≈80-residue motif, which occurs in many other RNA-binding proteins, contains two highly conserved regions (RNP1 and RNP2) that allow the motif to be recognized in newly sequenced proteins. X-ray crystallographic analysis has shown that the RNP motif consists of a four-stranded β sheet flanked on one side by two α helices. The conserved RNP1 and RNP2 sequences lie side by side on the two central β strands, and their side chains make multiple contacts with a single-stranded region of RNA. The single-stranded RNA loop lies across the surface of the β sheet and fits into a groove between the protein loop connecting strands β_2 and β_3 and the C-terminal region (Figure 11-10).

The *RGG box,* another RNA-binding motif found in hnRNP proteins, contains five Arg-Gly-Gly (RGG) repeats with several interspersed aromatic amino acids. Although the structure of this motif has not yet been

- Overview of mRNA processing in...
- Figure 11-8. Capping of the 5′ end...
- Figure 11-9. Visualization of hnRNP protein associated...
- Figure 11-10. Structure of complex between an...
- Figure 11-11. Hybridization of RNA molecules in...
- Figure 11-12. Model for cleavage and polyadenylation...
- Figure 11-13. Demonstration that introns are spliced...
- Figure 11-14. Consensus sequences around 5′ and...
- Figure 11-15. Analysis of RNA products formed...
- Figure 11-16. Splicing of exons in pre-mRNA...
- Figure 11-17. Diagram of interactions between pre-mRNA,...
- Figure 11-18. Electron micrograph of a spliceosome.
- Figure 11-19. The

determined, its arginine-rich nature is similar to the RNA-binding domains of the λ-phage N and HIV Tat proteins.

The 45-residue *KH motif* is found in the hnRNP K protein and several other RNA-binding proteins; commonly two or more copies of the KH motif are interspersed with RGG repeats. The three-dimensional structure of a representative KH motif, determined by NMR methods (Section 3.5), is similar to that of the RNP motif but smaller, consisting of a three-stranded β sheet supported from one side by a single α helix. It is not yet clear how this motif binds RNA. Mutations in the fragile-X gene *(FMR1),* which encodes a protein containing the KH motif, are associated with the most common form of heritable mental retardation. Although the molecular function of the Fmr1 protein is unknown, it presumably involves RNA binding. TOP

hnRNP Proteins May Assist in Processing and Transport of mRNAs

The association of pre-mRNAs with hnRNP proteins may prevent formation of short secondary structures dependent on base-pairing of complementary regions, thereby making the pre-mRNAs accessible for interaction with other macromolecules (Figure 11-11). Moreover, pre-mRNAs associated with hnRNP proteins present a more uniform substrate for further processing steps than would free, unbound pre-mRNAs each type of which forms a unique secondary structure dependent on its specific sequence.

The diversity of hnRNP proteins suggests that they probably have other functions as well. For example, various hnRNP proteins may interact with the RNA sequences that specify RNA splicing or cleavage/polyadenylation and contribute to the structure recognized by RNA-processing factors. Finally, cell-fusion experiments have shown that some hnRNP proteins remain localized in the nucleus, whereas others cycle in and out of the cytoplasm, suggesting that they function in the transport of mRNA (see later section). TOP

Pre-mRNAs Are Cleaved at Specific 3′ Sites and Rapidly Polyadenylated

In animal cells, all mRNAs, except histone mRNAs, have a 3′ poly(A) tail. Early studies of pulse-labeled adenovirus and SV40 RNA demonstrated that the viral primary transcripts extend beyond the poly(A) site in the viral mRNAs. These results suggested that A residues are added to a 3′ hydroxyl generated by endonucleolytic cleavage, but the predicted downstream RNA fragments are degraded so rapidly in vivo that they cannot be detected. However, this cleavage mechanism was firmly established by detection of both predicted cleavage products in in vitro processing reactions performed with extracts of HeLa-cell nuclei.

Early sequencing of cDNA clones from animal cells showed that nearly all mRNAs contain the sequence AAUAAA 10 – 35 nucleotides upstream from the poly(A) tail. Polyadenylation of RNA transcripts from transfected

spliceosomal splicing cycle.
Figure 11-20. Schematic diagrams comparing the secondary...
Figure 11-21. Localization of polyadenylated RNA and...
Figure 11-22. Transmission electron micrograph showing the...

genes is virtually eliminated when template DNA encoding the AAUAAA sequence is mutated to any other sequence except one encoding AUUAAA. The unprocessed RNA transcripts produced from such mutant templates do not accumulate in nuclei, but are rapidly degraded. Further mutagenesis of sequences within a few hundred bases of poly(A) sites revealed that a second signal downstream from the cleavage site is required for efficient **cleavage and polyadenylation** of most pre-mRNAs in animal cells. This downstream poly(A) signal is not a specific sequence but rather a GU-rich or simply a U-rich region within ≈50 nucleotides of the cleavage site.

Identification and purification of the proteins required for cleavage and polyadenylation of pre-mRNA has led to the model shown in Figure 11-12. According to this model, a 360-kDa *cleavage and polyadenylation specificity factor* (CPSF), composed of four different polypeptides, first forms an unstable complex with the upstream AU-rich poly(A) signal. Then at least three additional proteins — a 200-kDa heterotrimer called *cleavage stimulatory factor* (CStF), a 150-kDa heterotrimer called *cleavage factor I* (CFI), and a second cleavage factor (CFII), as-yet poorly characterized — bind to the CPSF-RNA complex. Interaction between CStF and the GU- or U-rich downstream poly(A) signal stabilizes the multiprotein complex. Finally, a *poly(A) polymerase* (PAP) binds to the complex before cleavage can occur. This requirement for PAP binding links cleavage and polyadenylation, so that the free 3′ ends generated are rapidly polyadenylated. Assembly of this large, multiprotein cleavage-polyadenylation complex around the AU-rich poly(A) signal in a pre-mRNA is analogous in many ways to formation of the transcription-initiation complex at the AT-rich TATA box of a template DNA molecule (see Figure 10-50). In both cases, multiprotein complexes assemble cooperatively through a network of specific protein – nucleic acid and protein-protein interactions.

Following cleavage at the poly(A) site, polyadenylation proceeds in two phases. Addition of the first 12 or so A residues occurs slowly, followed by rapid addition of up to 200 – 250 more A residues. The rapid phase requires the binding of multiple copies of a poly(A)-binding protein containing the RNP motif. This protein is designated *PABII* to distinguish it from the poly(A)-binding protein that binds to the poly(A) tail of cytoplasmic mRNAs. PABII binds to the short A tail initially added by PAP, stimulating polymerization of additional A residues by PAP (see Figure 11-12). PABII is also responsible for signaling poly(A) polymerase to terminate polymerization when the poly(A) tail reaches a length of 200 – 250 residues, although the mechanism for measuring this length is not yet understood. ↑ TOP

Splicing Occurs at Short, Conserved Sequences in Pre-mRNAs via Two Transesterification Reactions

During the final step in formation of a mature, functional mRNA, the introns are removed and exons are spliced together (see Figure 11-7). The discovery that introns are removed during splicing came from electron

MOLECULAR CELL BIOLOGY
Lodish Berk Zipursky Matsudaira Baltimore Darnell
W H FREEMAN AND COMPANY

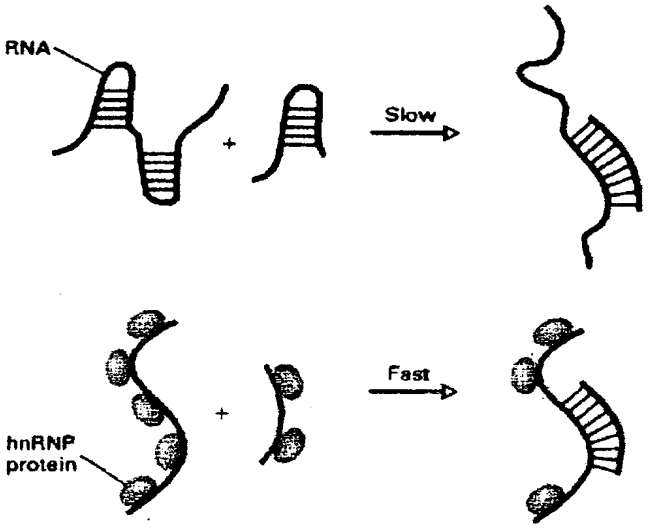


Figure 11-11. Hybridization of RNA molecules in vitro is accelerated by hnRNP proteins. The presence of complex secondary structures within RNA molecules inhibits hybridization between long complementary sequences in separate molecules. Association of hnRNP proteins with RNA is thought to prevent formation of RNA secondary structures, thereby facilitating base-pairing between different complementary molecules. These proteins may have a similar function in vivo. [Adapted from D. S. Portman and G. Dreyfuss, 1994, *EMBO J.* **13**:213.]